

Is the observed persistence spurious? A test for fractional integration versus short memory and structural breaks.*

Laura Mayoral

Department of Economics and Business

Universitat Pompeu Fabra.

Abstract

Although it is commonly accepted that most macroeconomic variables are non-stationary, it is often difficult to identify the source of the non-stationarity. In particular, it is well-known that integrated and short memory models containing trending components that may display sudden changes in their parameters share some statistical properties that make their identification a hard task. The goal of this paper is to extend the classical testing framework for $I(1)$ versus $I(0)+$ breaks by considering a more general class of models under the null hypothesis: non-stationary fractionally integrated (FI) processes. A similar identification problem holds in this broader setting which is shown to be a relevant issue from both a statistical and an economic perspective. The proposed test is developed in the time domain and is very simple to compute. The asymptotic properties of the new technique are derived and it is shown by simulation that it is very well-behaved in finite samples. To illustrate the usefulness of the proposed technique, an application using inflation data is also provided.

*I would like to thank Jushan Bai and Manuel Domínguez for his very useful comments. I acknowledge financial support from the Spanish Ministry of Education through grants SEC2003-04429 and SEC2003-04476 and also from the Barcelona Economics Program of CREA.

1. INTRODUCTION

Most macroeconomic variables seem to display a non-stationary and very persistent behavior (Granger, 1966). These variables are typically represented by models that are composed of a trend and a cycle. Before Nelson and Plosser (1982), it was commonplace to assume that the trend was linear. Nevertheless they showed that the hypothesis of a random walk in the trend component could not be rejected for many widely used aggregate macroeconomic variables and then, the use of unit roots became very popular. Processes containing a unit root are very persistent, to the extent that all shocks have a permanent effect on the future dynamics of the process.

This approach encountered soon many counterchallenges since the latter hypothesis is hardly plausible for many economic variables. One of the most constructive was suggested by Perron (1989) who proposed to consider models in which most shocks vanished quite fast while only those related to very significant events (such as wars, deep economic crisis, etc.) had a permanent impact. These ideas could be captured by using trend components that might display sudden changes in their parameters (structural breaks) in an otherwise short-memory process. Furthermore, he showed that breaks in the trend produce serial correlation patterns that are similar to those of an integrated process. As a consequence, classical unit root tests are not able to reject the latter hypothesis when in fact the DGP is a short memory+breaks process, in spite of the very different dynamics and implications that these models have. See Perron (2005) for a comprehensive survey on this area.

Another alternative to the unit root framework was suggested by Hosking (1981) and Granger and Joyeux (1980) who, by considering fractional orders of integration showed that it was possible to obtain processes with a richer class of persistence properties and long-run behaviours. More specifically, fractionally integrated (FI) processes can accommodate long memory (stationarity with hyperbolic correlation decay), non-stationary mean-reversion (very persistent non-stationary processes with non-permanent shocks) and non-stationarity without mean reversion (for instance, unit root processes). There is wide empirical and theoretical support for the existence of these features in both macroeconomic and financial

data that has motivated an active research in this field (see Henry and Zaffaroni, (2002) and Robinson (2004) for recent surveys on this area).

Not surprisingly, it is also difficult to provide an unambiguous answer as to whether a process is fractionally integrated or is short memory plus some deterministic components that may be perturbed by sudden changes, since the same identification problem as that of the $I(1)$ case holds here. The issue of detecting correlation patterns similar to those of a FI process when the DGP is short memory containing deterministic terms that may display breaks has been widely analyzed (see Mikosh and Starika, (2004), Bhattacharya et al. (1983), Giraitis et al. (2001), Perron (1989), Künsh (1986), Teverosky and Taquq (1997), Diebold and Inhoue (2001), Perron and Qu (2004) among many others). It is generally concluded that the use of standard techniques devised for *FI* processes could lead to the detection of spurious persistence when applied on short memory processes containing breaks. The opposite effect is also well-documented, that is, conventional procedures for detecting and dating structural changes tend to find spurious breaks, usually in the middle of the sample, when in fact there is only fractional integration in the data (see Nunes et al. (1995), Krämer and Sibbertsen (2002) and Hsu (2001)).

Consequently, although there is a broad consensus on the non-stationarity of most macro-variables, it is often difficult to be sure about the source of the non-stationarity, that is, whether it comes from a high degree of persistence or from the existence of parameter instability. From a statistical point of view, to solve correctly this issue is critical in order to avoid misspecifications. But the question is also well-motivated from an economic point of view. On the one hand, if a model is unstable, it would not be useful for policy simulations since the Lucas' critique (1976) would apply with full force. On the other, to have an accurate picture of the duration of shocks is crucial in order to design most economic policies. This has motivated a growing literature aimed to determine the persistence and stability of some key economic variables in both macroeconomics and finance that has given rise to interesting controversies¹ since opposite conclusions are often reached. Typically, these

¹To mention some examples, some authors have found strong persistence in series such as inflation, asset returns, GNP, etc., (see, among others, Pivetta and Reis (2004), Ding, Granger and Engle (1993), Lobato and

contributions only consider tests of integer integration versus short memory+structural breaks (even in cases where there is empirical support for the hypothesis of FI), the reason being that no formal test of non-stationary FI versus short memory+breaks is available. Then, if the model is in fact FI , contradicting results are likely to be found when different techniques are employed.²

The goal of this paper is to develop a simple testing device in the time domain that is able to determine whether the observed non-stationarity is due to fractional integration (that is able to encompass the $I(1)$ model as a particular case) or to the existence of deterministic trends, possibly containing breaks, in an otherwise short memory process.

To illustrate the simplicity and empirical usefulness of this technique, an application using inflation data has been considered. This variable offers a good example of the above-described problems, since in spite of the myriad of papers dealing with the study of its persistence and stability properties, no consensus has been reached yet. Economic and statistical support for the existence of FI behavior in inflation is provided and it is shown that under the latter hypothesis some of the contradicting results that have been recently found in the literature could be accounted for if inflation is FI .

In spite of the large literature for testing the hypotheses of $I(1)$ versus $FI(d)$ or versus $I(0)$ plus structural breaks, (SB), there are very few techniques that allow for testing FI vs. SB. To the best of our knowledge, this is the first one that allows for directly testing non-stationary FI vs. $I(0)$ +breaks. Other interesting contributions in this area are that of Sibbertsen and Venetis (2004) who proposed a technique based on the comparison of standard log-periodogram regression estimation of the memory parameter with its tapered counterpart but this method is only valid for processes with $d < 0.5$. Hidalgo and Robinson (1996) analyze the related, although different, problem of testing for structural breaks in a

Savin, (1998) and a long etc.). Nevertheless, these findings are often attributed to the existence parameter unstability (see Levin and Piger (2003), Mikosh and Starica (2004), etc.), therefore arriving to opposite conclusions about the duration of shocks to these variables.

²More specifically, tests of $I(1)$ vs. $I(0)$ + breaks would tend to reject the former hypothesis for large T if the DGP is $FI(d)$ with $d < 1$, but the power will be low in finite samples. On the other hand, as discussed above, tests of $I(0)$ vs. $I(0)$ + structural breaks will tend to find spurious breaks if the process is FI .

regression model when the residuals may exhibit long range dependence. Lazarova (2004) extends the latter framework by allowing for long memory also in the regressors. Finally, Dolado, Gonzalo and Mayoral (2005) propose an extension of the Perron (1989) and Zivot and Andrews (1992) approaches (developed for testing $I(1)$ vs. $I(0)$ plus breaks with known or unknown break date, respectively) along the lines of the Fractional Dickey-Fuller test (Dolado, Gonzalo and Mayoral, 2002). The properties of this technique are still under investigation.

The structure of the paper is the following. Section 2 presents the model and the hypotheses of interest. For the sake of clarity, the new test is presented in a sequence of steps: Section 3 analyzes the simpler problem of testing for FI vs $I(0)$ where no breaks are allowed to exist. This framework constitutes the basis of the test presented in Section 4 that allows for the presence of a break occurring at an unknown time. Asymptotic results as well as finite sample simulations (Section 5) are provided. To illustrate the usefulness and simplicity of the new technique, Section 6 analyzes the statistical properties of US inflation data. Section 7 draws some final conclusions. All proofs are gathered in Appendix A while critical values for the proposed tests are presented in Appendix B.

In the sequel, the definition of a $FI(d)$ process that we will adopt is that of an (asymptotically) stationary process, when $d < 0.5$ and that of a non-stationary (truncated) process, when $d \geq 0.5$. Those definitions are similar to those used in, e.g., Robinson (1994) or Tanaka (1999) (see, Appendix A in Dolado, Gonzalo and Mayoral (2002) for details). Moreover, the following conventional notation is adopted throughout the paper: $[\cdot]$ indicates integer part, L is the lag operator, $\Delta = (1 - L)$, $\Gamma(\cdot)$ denotes the gamma function, $\{\pi_i(d)\}$ represents the sequence of coefficients associated to the expansion of Δ^d in powers of L and are defined as

$$\pi_i(d) = \frac{\Gamma(i-d)}{\Gamma(-d)\Gamma(i+1)}. \quad (1)$$

All integrals are taken with respect to the Lebesgue measure; $B_d(\cdot)$ is standard fractional Brownian motion (fBM) corresponding to the limit distribution of the standardized partial sums asymptotically stationary (truncated) $FI(d)$ processes³; Finally, \xrightarrow{w} and \xrightarrow{p} denote

³According to the notation introduced in Marinucci and Robinson (1999) $B_d(\cdot)$ is a type II Brownian

weak convergence and convergence in probability, respectively.

2. THE MODEL AND THE HYPOTHESES

In the following it is assumed that the data y_1, \dots, y_T is generated as,

$$y_t = \beta' Z_t + \delta' V_t(\omega) + x_t, \quad t = 1, 2, \dots, \quad (2)$$

where,

$$\Delta^d x_t = u_t, \quad x_t = 0 \text{ for all } t \leq 0, \quad (3)$$

and,

$$V_t(\omega) = \begin{cases} Z_{t-T_B} & t \geq T_B, \\ 0 & \text{otherwise.} \end{cases}$$

The process y_t and the $k \times 1$ vector Z_t of non-stochastic variables⁴ are observable, β and δ are $k \times 1$ vectors or parameters and $\{u_t\}$ is an unobserved zero-mean process whose spectral density is strictly positive at zero frequency. Moreover, $\{u_t\}$ is assumed to have a Wold representation $u_t = \Psi(L) \varepsilon_t$, where the coefficients ψ_j are such that $\sum_{j=0}^{\infty} j |\psi_j| < \infty$ and $\{\varepsilon_t\}$ is an unobserved *i.i.d.* zero mean process with unknown variance equal to σ^2 and $\mu_4 = E|\varepsilon|^4 = \eta\sigma^4 < \infty$. T_B is the parameter describing the time when the break, if it exists, occurs and the parameter $\omega = T_B/T$ determines the location of the break point in the sample. It verifies that

$$\omega \in \Omega = [\omega_L, \omega_H] \subset (0, 1).$$

Recall that the objective of this paper is to determine the source of the non-stationarity observed in a data set, more specifically whether it comes from a high degree of inertia or from parameter unstability. Hence, the suspected non-stationarity is modelled in two different ways. Under the null hypothesis y_t is consider to be a non-stationary $FI(d_0)$ process with no breaks. Then, δ is assumed to be equal to zero (no breaks) and non-stationarity of y_t requires $d \geq 0.5$.⁵ Under H_1 , y_t is short memory ($d_0 = 0$), but δ is

motion

⁴ Z_t will typically contain polynomials in t .

⁵ For any $d > 0$, y_t is FI and is stationary and invertible as long as $-0.5 < d < 0.5$.

(partially or totally) unrestricted, allowing in this way for the possibility of breaks. It will not be needed to make additional parametric assumptions on the structure of u_t but, if u_t admits an ARMA(p,q) representation, y_t will be an ARFIMA (p, d, q) process, (see Hosking, (1981) and Granger and Joyeux, (1980)) under H_0 or a trend-stationary ARMA(p,q) model containing breaks under the alternative.

Therefore, if y_t is defined as in (2), the null and the alternative hypotheses can be written as,

$$H_0 : d = d_0, \delta = 0 \text{ for a } d_0 > 0.5,$$

and,

$$H_1 : d = 0,$$

where at least some of the components of δ are unrestricted.

3. PRELIMINARIES: POINT OPTIMAL TESTS FOR FRACTIONAL PROCESSES

The aim of this section is to present the main ideas that later on (Section 4) will be used to build the test for $FI(d)$ vs. $I(0) + breaks$. In particular, this section introduces as a preliminary step a procedure for testing $FI(d_0)$ versus $FI(d_1)$ when no breaks in the data are suspected. For the sake of clarity, the following subsection discusses the simplest case where no deterministic components are allowed for, while subsection 3.2 extends this framework by including such components.

3.1 Model without deterministic components

Let us first consider the DGP defined in (2) with $\beta = \delta = 0$ under both hypotheses.

Although the goal of this section is to develop a test for (non-stationary) FI vs. short memory ($I(0)$), under the restriction $\beta = \delta = 0$ it is easy to consider a slightly more general framework. To illustrate how this generalization could be accomplished, in this subsection the integration order of y_t under the alternative hypothesis, d_1 , would be allowed to be in the interval $0 \leq d_1 < d_0 < 1.5$. Therefore, under H_1 , y_t could be short memory ($d_1 = 0$) but

also long memory ($d_1 \in (0, 0.5)$) and even non-stationarity ($d_1 > 0.5$). Also, the integration order of y_t under the null, d_0 , would not be restricted to the non-stationary range of values of d , but it could be any value bigger than d_1 . Then, the *DGP* will be,

$$y_t = x_t \quad (4)$$

$$\Delta^{d_i} x_t = u_t, \quad i = \{0, 1\}.$$

Under the null hypothesis, the integration order of x_t is d_0 whereas under the alternative hypothesis is $d_1 \in [0, d_0)$. The problem of testing $FI(d_0)$ vs. $FI(d_1)$ can be seen as a simple hypotheses test and, as such, the Neyman-Pearson Lemma provides for the most powerful test.⁶ The power of this test would be an upper bound for the power function of any test based on the same likelihood. Under gaussianity, minus two times the log-likelihood function under the null hypothesis is (except for an additive constant) given by,

$$L(d, \sigma)|_{H_0} = \left(\Delta^{d_0} y \right)' \Sigma^{-1} \left(\Delta^{d_0} y \right),$$

where $\Delta^{d_0} y = (\Delta^{d_0} y_1, \dots, \Delta^{d_0} y_T)'$ and Σ is the non-singular variance-covariance matrix for (u_1, \dots, u_T) . Analogously, under the alternative hypothesis of $FI(d_1)$, it is equal to,

$$L(d, \sigma)|_{H_1} = \left(\Delta^{d_1} y \right)' \Sigma^{-1} \left(\Delta^{d_1} y \right),$$

with $\Delta^{d_1} y = (\Delta^{d_1} y_1, \dots, \Delta^{d_1} y_T)'$. By the Neyman-Pearson Lemma, the most powerful test of the null hypothesis of $d = d_0$ vs. $d = d_1$ rejects the null hypothesis for small values of the likelihood ratio statistic $L(d, \sigma)|_{H_1} - L(d, \sigma)|_{H_0}$. Consider first the simplest case where $u_t = \varepsilon_t$, so that $\Sigma = \sigma^2 I_T$. The critical region (CR) can be written as,

$$\frac{\Delta^{d_1} y' \Delta^{d_1} y - \Delta^{d_0} y' \Delta^{d_0} y}{\sigma^2} < k_T$$

Since σ^2 is unknown, it should be replaced by a consistent estimator. Under the null hypothesis, the obvious choice would be $T^{-1} (\Delta^{d_0} y' \Delta^{d_0} y)$. After some manipulation, it is obtained that the UMP rejects the null hypothesis for small values of the statistic,

⁶This approach is similar to the one proposed by King (1988), Dufour and King (1991) and Elliott, Rothemberg and Stock (1996) who analyzed Neyman-Pearson's type tests in the context of testing for unit roots against local autoregressive alternatives.

$$R = \frac{\Delta^{d_1} y' \Delta^{d_1} y}{\Delta^{d_0} y' \Delta^{d_0} y}. \quad (5)$$

The statistic in (5) is similar to the Von-Neumann ratio proposed in the framework of efficient unit root tests (see Sargan and Bhargava (1983) and Bhargava (1986)). These authors considered the problem of testing for a random walk versus an AR(1) process. To do that, they proposed as test statistic a ratio of variances similar to (5) and they showed that this statistic was locally most powerful.⁷ Notice that their statistic only has this Neyman-Pearson interpretation when the value of the autoregressive parameter is equal to zero under the alternative hypothesis.

It is possible to calculate the exact null distribution of R under Gaussianity via Imhof's algorithm, since R involves a ratio of quadratic forms. Also, an approximate finite-sample optimality theory might be constructed along the lines of that of Bhargava (1986). The following theorem describes the asymptotic distribution of the proposed test statistic in a more general situation where the assumption of gaussianity is not needed.

Theorem 1 *Let y_t be defined as in (4) with $u_t = \varepsilon_t$ and $0 \leq d_1 < d_0 < 1.5$. Then, under the null hypothesis of $FI(d_0)$, with $d_0 < 1.5$ the asymptotic distribution of (5) is given by,*

1. *if $(d_0 - d_1) > 0.5$,*

$$T^{1-2(d_0-d_1)} \frac{\Delta^{d_1} y' \Delta^{d_1} y}{\Delta^{d_0} y' \Delta^{d_0} y} \xrightarrow{w} \int_0^1 B_{(d_0-d_1)}^2(r) dr, \quad (6)$$

2. *and if $(d_0 - d_1) < 0.5$*

$$\begin{aligned} T^{1/2} \left(\frac{\Delta^{d_1} y' \Delta^{d_1} y}{\Delta^{d_0} y' \Delta^{d_0} y} - \frac{\Gamma(1-2(d_0-d_1))}{\Gamma^2(1-(d_0-d_1))} \right) &\xrightarrow{w} N(0, V), \text{ if } (d_0 - d_1) < 1/4 \\ T^{1-2(d_0-d_1)} \left(\frac{\Delta^{d_1} y' \Delta^{d_1} y}{\Delta^{d_0} y' \Delta^{d_0} y} - \frac{\Gamma(1-2(d_0-d_1))}{\Gamma^2(1-(d_0-d_1))} \right) &\xrightarrow{w} RS, \text{ if } 1/4 < (d_0 - d_1) < 1/2 \end{aligned} \quad (7)$$

where $B_\delta(r)$ is a fBM as defined in Marinucci and Robinson (1999) and RS is the Roseblatt distribution. Moreover, under gaussianity this is the most powerful test.

⁷Schmidt and Phillips (1992) showed that for a Gaussian likelihood the Lagrange multiplier principle also leads to this expression.

The results above can be generalized to values of d_0 greater than 1.5, following Ming (1998). Notice that the distribution only depends on the distance between the hypotheses, $(d_0 - d_1)$, and not on the particular value of d_0 .⁸ Whenever $(d_0 - d_1 > 0.5)$, the asymptotic distribution is a functional of fractional Brownian motions. This is so because the process in the numerator, $\Delta^{d_1} y_t$, is $FI(d_0 - d_1)$, with $(d_0 - d_1) > 0.5$ and therefore, non-stationary. Well-known results guarantee the convergence to fBM in this case. The situation is more complicated when the filtered process $\Delta^{d_1} y_t$ is stationary, that is, whenever $(d_0 - d_1 < 0.5)$. Even in the very simple framework considered in this subsection, two different asymptotic distributions arise according to the values of the difference $(d_0 - d_1)$. The asymptotic distribution will be normal as long as the difference $(d_0 - d_1)$ is smaller than $1/4$. Otherwise, the Roseblatt distribution applies (see Hosking, (1996)).⁹ Finally, notice that the distribution depends upon the value of d_1 and therefore no uniform most powerful test exists.

The following theorem states the consistency of the test proposed above. It turns out that it remains consistent as long the true integration order is smaller than the value employed as null hypothesis when running the test.

Theorem 2 *Let y_t be a $FI(d^*)$ process defined as in (4) with $u_t = \varepsilon_t$. Then, the probability of rejecting the null hypothesis of $d = d_0$ tends to 1 as long as $d_0 > d^*$.*

This result is very interesting because it underlines the importance of explicitly considering fractional alternatives and not just integer ones. To see this more clearly, suppose that the true integration order is $d^* = 0.7$ but a test of $I(1)$ vs. $I(0)$ is implemented. According to the result of Theorem 2, since $d_0 = 1 > d^* = 0.7$, the test would tend to reject the $I(1)$ hypothesis in large samples. This suggests that traditional methods for testing $I(1)$ vs. $I(0)$ (with or without structural breaks) could be overrejecting the hypothesis of strong persistence in favor of short memory when the true model is fact strong persistent but with

⁸The cases $(d_0 - d_1) = \{1/2, 1/4\}$ are discontinuity points in the asymptotic theory and they are not considered in this theorem.

⁹Asymptotic distributions that depend upon the distance between the hypotheses have also been found in Dolado, Gonzalo and Mayoral (2002) in a similar context.

an integration order less than one. Section 5 reports a simulation study confirming the arguments above (see Table 5.4).

Clearly, in the context of Theorem 1 d_1 is not a single alternative but belongs to a set of values. A complete description of this case, then, would entail the analysis of the power envelope and the local asymptotic power, along the lines of Elliott et al. (1996) or Dufour and King (1991). Nevertheless, this analysis will be skipped here for the sake of brevity, since this is not the main goal of the paper. As it was stated in the introduction, the interest of this paper is to test $FI(d)$ against alternatives that are short memory (possibly with deterministic trends and breaks). This implies that the alternative hypothesis needs to be $d_1 = 0$. Therefore, our framework is different from that of the above-mentioned papers in the sense that in this case H_1 is in fact a single point. Henceforth only the case $d_1 = 0$ will be considered.

Theorems 1 and 2 just deal with the simplest case where $\Delta^{d_0}y_t = \varepsilon_t$ does not have short run structure. In the more general situation where $\Delta^{d_0}y_t = u_t$, and u_t is a short memory process, the distributions in Theorem 1 would depend on some nuisance parameters and therefore, should be corrected. If the parametric form of u_t is known, an estimate of the variance-covariance matrix Σ can be obtained and therefore the critical region of the most powerful test would be given by,

$$\frac{(\Delta^{d_1}y)' \hat{\Sigma}_1^{-1} (\Delta^{d_1}y)}{(\Delta^{d_0}y)' \hat{\Sigma}_0^{-1} (\Delta^{d_0}y)} < k_T,$$

for consistent estimators, $\hat{\Sigma}_0$ and $\hat{\Sigma}_1$ of Σ , under H_0 and H_1 respectively. In the general case where no information of the parametric form of u_t is available, the statistic in (5) can still be employed as long as a correction that takes into account the correlation structure of u_t is introduced. In the following, we explore in more detail the second alternative since it has the advantage that no more additional assumptions on the structure of u_t need to be adopted. The following theorem presents the asymptotic behavior of the statistic defined in (5) for the case where u_t verifies the assumptions of Section 2. As mentioned before, attention will be restricted to the case where $d_1 = 0$ and $d_0 > 0.5$.

Theorem 3 *Let y_t be defined as in (4) with $d_1 = 0$ and d_0 is a value greater than 0.5.*

Then, under the null hypothesis of $FI(d_0)$ the asymptotic distribution of (5) is given by:

$$T^{1-2d_0} \frac{y'y}{\Delta^{d_0} y' \Delta^{d_0} y} \xrightarrow{w} \frac{\lambda^2}{\gamma_0} \int_0^1 B_{d_0}^2(r) dr,$$

where $\lambda = \sigma \Psi(1)$ and $\gamma_0 = \sigma^2 \sum_{j=0}^{\infty} \psi_j^2$.

The existence of correlation structure in u_t introduces the nuisance term λ^2/γ_0 in the asymptotic distribution. It is possible to estimate this factor by nonparametric kernel techniques, analogous to those that are used in the estimation of the spectral density (see Andrews (1991)). Then, the statistic in (5) can be corrected to account for the short term correlation so that it is possible to recover the same asymptotic distributions as in Theorem 1. In particular, γ_0 can be estimated under the null hypothesis simply by $\sum (\Delta^{d_0} y_t)^2 / T$ whereas λ^2 can be rewritten as:

$$\lambda^2 = \gamma_0 + 2 \sum_{i=1}^{\infty} \gamma_i = 2\pi s_u(0).$$

Several estimators of this quantity have been proposed, see Andrews (1991) for an analysis and comparison of the different techniques. One of the most popular is the Newey-West estimator:

$$\hat{\lambda}^2 = \hat{\gamma}_0 + 2 \sum_{i=1}^q (1 - j/(q+1)) \hat{\gamma}_i$$

where $\hat{\gamma}_i = T^{-1} \sum_{t=j+1}^T u_t u_{t-j}$. Andrews (1991) also provides a guideline for choosing the value of the lag truncation, q .¹⁰

3.2. Unknown deterministic components

Since most macroeconomic time series usually display a trending behavior or a level different from zero, the test presented above need to be adapted in order to be invariant to these components. Consider again model (2) where, in contrast to the previous subsection, β is a vector of unknown constants. The existence of breaks is not allowed until Section

¹⁰As Andrews points out, a correct choice of q is very important since the performance of these estimators can greatly depend on this choice.

4, so again, δ is assumed to be equal to zero under both H_0 and H_1 . When β is unknown, the process x_t in (2) is unobservable and then, an estimate of β is needed to carry out the test. It follows from the Grenander-Rosenblatt theorem (Grenander and Rosenblatt, 1957) that if a process is (trend) stationary, the trend function can be efficiently estimated by an *OLS* regression. Then, taking the appropriate differences under both the null and under the alternative hypothesis, a simple *OLS* regression would yield efficient estimates of β . Equivalently, these estimates can be seen as the results of two constrained GLS regressions, one imposing $d = d_0$ and the other imposing $d = 0$. It follows (Lehmann (1959)) that the most powerful invariant (MPI) test will reject the null hypothesis of $d = d_0$ for small values of $\min_{\beta} L(d = 0, \sigma, \beta) - \min_{\beta} L(d = d_0, \sigma, \beta)$. Then, the MPI test will reject the null hypothesis for small values of the expression,

$$\min_{\beta} (y - Z\beta)' \Sigma^{-1} (y - Z\beta) - \min_{\beta} \left(\Delta^{d_0} y - \Delta^{d_0} Z\beta \right)' \Sigma^{-1} \left(\Delta^{d_0} y - \Delta^{d_0} Z\beta \right), \quad (8)$$

with $Z = (Z'_1, \dots, Z'_T)'$.

As it is common in this literature, the asymptotic behavior of the statistics are quite sensitive to the nature of the regressors included in Z_t . In the following, we present explicit formulas for the constant mean ($Z_t = 1$) and the linear time trend case ($Z_t = (1, t)$). The critical regions of the MPI tests (in the simplest case where $u_t = \varepsilon_t$ and $d_1 = 0$) can be written, after rearranging terms, as

$$R^c = \frac{\sum_{t=1}^T (y_t - \hat{\alpha}_1)^2}{\sum_{t=2}^T (\Delta^{d_0} (y_t - \hat{\alpha}_0))^2} < k'_T \text{ and } R^r = \frac{\sum_{t=1}^T (y_t - \hat{\alpha}_1 - \hat{\beta}_1 t)^2}{\sum_{t=2}^T (\Delta^{d_0} (y_t - \hat{\alpha}_0 - \hat{\beta}_0 t))^2} < k'_T, \quad (9)$$

for the mean and mean+trend respectively, where $\hat{\alpha}_i$ and $\hat{\beta}_i$ are the OLS estimators in the constrained models. Notice that under H_0 , the process should be differenced prior to the estimation of the deterministic components in order to have short memory residuals. Then, $\hat{\alpha}_0$ and $\hat{\beta}_0$ are computed as the OLS estimates in a regression of $\Delta^{d_0} y_t$ on $\Delta^{d_0} 1_{(t>0)}$ and $\Delta^{d_0-1} 1_{(t>0)}$ (or equivalently, $\Delta^{d_0} t$), where $\Delta^{\delta} 1_{(t>0)} = \sum_{i=0}^{t-1} \pi_i(\delta)$ and the coefficients $\pi_i(\cdot)$ are defined in (1). Under H_1 , in turn, $\hat{\alpha}_1$ and $\hat{\beta}_1$ are simply the OLS estimates of the original process, y_t on a constant or a constant and a trend for the R^c and R^r , respectively.

The following theorem presents the asymptotic properties of the test for the general case where u_t is allowed to have serial correlation and a constant term or a constant and a linear trend are included under both hypotheses. Notice that this framework contains as a particular case the Sargan and Bhargava statistic for testing $d_0 = 1$ and $d_1 = 0$.

Theorem 4 *Let y_t defined as in (2) with $\delta = 0$ under both H_0 and H_1 . Then, under the null hypothesis of $FI(d_0)$, with $d_0 \geq 0.5$, and $Z_t = (1)$ or $Z_t = \begin{pmatrix} 1 & t \end{pmatrix}$, the asymptotic distributions of R^c and R^r are given, respectively, by*

- If $Z_t = (1)$,

$$T^{1-2d_0} R^c = T^{1-2d_0} \frac{\sum (y_t - \hat{\alpha}_1)^2}{\sum (\Delta^{d_0}(y_t - \hat{\alpha}_1))^2} \xrightarrow{w} \frac{\lambda^2}{\gamma_0} \int_0^1 B_{d_0}^{\mu 2}(r) dr. \quad (10)$$

where $B_{d_0}^\mu$ is a “demeaned” fBM, see Appendix A for its exact form.

- and if $Z_t = \begin{pmatrix} 1 & t \end{pmatrix}$,

$$T^{1-2d_0} R^r = T^{1-2d_0} \frac{\sum (y_t - \hat{\alpha}_1 - \hat{\beta}_1 t)^2}{\sum (\Delta^{d_0}(y_t - \hat{\alpha}_1 - \hat{\beta}_1 t))^2} \xrightarrow{w} \frac{\lambda^2}{\gamma_0} \int_0^1 B_{d_0}^{\tau 2}(r) dr.$$

where $B_{d_0}^\tau$ is a “detrended” fBM, see Appendix A for its exact form.

As in the $I(1)$ vs. $I(0)$ case, the estimation of deterministic components has a non-negligible asymptotic effect. Critical values for the demeaned and detrended distributions are reported in Appendix B, corresponding to the case where $u_t = \varepsilon_t$. In the general case where autocorrelation in u_t is suspected, a semiparametric correction, as explained in subsection 3.1, should be computed.

4. TESTING FRACTIONAL INTEGRATION VERSUS STRUCTURAL BREAKS

In this section the assumption of $\delta = 0$ under H_1 is relaxed, allowing in this way for the existence of breaks under the latter hypothesis. Perron (1989) was one of the first

to show that standard unit root tests could conduct to misleading conclusions if the true *DGP* was an $I(0)$ process containing breaks in the deterministic components. This seminal contribution was the starting point of a myriad of articles on the problem of distinguishing between $I(1)$ vs. $I(0) + \text{breaks}$.

Surprisingly, there are very few contributions dealing with the more general topic of testing for $FI(d)$ vs. $I(0) + \text{breaks}$. Some authors have considered the problem of testing for stationary $FI(d)$ versus short memory with monotonic or non-monotonic trends (Kunsch (1986) and Sibbertsen and Venetis (2004), respectively). Another line of research has explored the related topic of how to test for breaks in a long memory context (see Hidalgo and Robinson (1996) and Lazaroza (1994)) but again, these contributions only consider stationary $FI(d)$ processes and therefore cannot be applied in the framework of this paper. Therefore, to the best of our knowledge, this is the first contribution offering a unified framework for testing non-stationary fractional integration (that includes the $I(1)$ model as a particular case) vs. trend stationarity with parameter instability.

As in Section 3, the focus is placed on processes that under H_0 are $FI(d_0)$, with $0.5 < d_0 < 1.5$. This interval contains the unit root case but also other interesting behaviors that have been shown of empirical relevance in economics. Under H_1 , the process of interest is characterized as being $I(0)$ with a possibly breaking trend. As in Perron (1989) and Zivot and Andrews (1992), several models are considered according to the included deterministic terms and the parameters that are allowed to break. For simplicity, attention will be restricted to the case where there exists at most a single break. An extension to a multiple-change environment can be entertained along the lines of Bai (1999) and Bai and Perron (1998).

Let y_t be defined as in (2). As discussed in Section 2, the null hypothesis is characterized by the joint hypothesis $d = d_0$ and $\delta = 0$. Then, under H_0 , y_t is the sum of a (fractionally) integrated component and some (smooth) deterministic terms. Alternatively, under H_1 , the process x_t is short memory and therefore $d = 0$. Breaks are allowed by letting δ unrestricted.

If the time when the break takes place was known, the matrix $V_t(\omega)$ would be completely determined and then, the test would reject the null hypothesis of (stable) fractional integra-

tion for small values of $\inf_{\beta, \delta} L(d_1 = 0, \Sigma, \beta, \delta) - \inf_{\beta} L(d = d_0, \Sigma, \beta, \delta = 0)$. In other words, the critical region would be given by,

$$\inf_{\beta, \delta} (y - Z\beta - V\delta)' \Sigma^{-1} (y - Z\beta - V\delta) - \inf_{\beta} \left(\Delta^{d_0} y - \Delta^{d_0} Z\beta \right)' \Sigma^{-1} \left(\Delta^{d_0} y - \Delta^{d_0} Z\beta \right) < k_T, \quad (11)$$

for some k_T . Moreover, under gaussianity this will be the MPI test. But, since the date break is in general unknown, the candidate for break point would be chosen as the one that maximizes the likelihood (or alternatively, that minimizes the variance). As a consequence of this estimation, optimality is lost. Two distinct cases may arise. The first is when the interest is centered on change points in a known restricted interval, say $\Omega = [\omega_L, \omega_H]$ for $0 < \omega_L < \omega_H < 1$. This would be the case when one wants to test for changes initiated by some institutional or political change that has occurred at a known time period. The second is the case where no information is available a priori and hence, all points in $(0, 1)$ are of some interest. This situation may arise when one wants to apply a test of structural break as a general diagnostic test of model adequacy. But considering the whole interval $(0, 1)$ would result in tests with very low power. Then, the minimization is carried out in $\omega \in \Omega$, where $\Omega = [\omega_L, \omega_H]$ for some $0 < \omega_L < \omega_H < 1$. More specifically, when no information on the location of the break is available, we will use the restricted interval $\Omega = [0.15, 0.85]$, following the suggestions in Andrews (1993). In the case where there is no short term correlation in u_t , the test statistic becomes,

$$R_b = \frac{\inf_{\omega \in \Omega} (\sum_{t=1}^T (y_t - \hat{\beta}' Z_t - \hat{\delta}(\omega)' V_t(\omega))^2)}{\sum_{t=2}^T (\Delta^{d_0} (y_t - \hat{\beta}' Z_t))^2} \quad (12)$$

The asymptotic distribution depends upon the regressors contained in Z_t and also on the parameters that are allowed to break. In the following, we will analyze the four cases considered by Perron (1989) and Zivot and Andrews (1992). In three of these models, Z_t contains both a constant and a linear trend but they differ on the parameters that are allowed to break: Model 1 allows for an exogenous break in the level of the series, Model 2 allows for a change in the rate of growth and finally, Model 3 admits both changes. In addition, we also consider “Model 0”¹¹, where Z_t only contains a constant that is allowed

¹¹This case may be of interest when modeling series who do not seem to display a trend, such as inflation

to break once in the sample.

To facilitate the statement of the theorem, the notation will be simplified by defining the following dummy variables: $DC_t = 1$, if $t > T_B$ and 0 otherwise and $DT_t = (t - T_B)$ if $t > T_B$ and 0 otherwise.

Theorem 5 *Let y_t defined as in (2) where Z contains a constant or a constant and a linear time trend. Then, under the null hypothesis of $FI(d_0)$, with $0.5 < d_0 < 1.5$, the asymptotic distribution of (12) is given by,*

- If $Z_t = (1)$,

$$\begin{aligned} Model 0 : R_b^0 &= T^{1-2d_0} \frac{\inf_{\omega \in \Omega} \sum (y_t - \hat{\alpha}_1 - (\hat{\alpha}_2 - \hat{\alpha}_1) DC_t)^2}{\sum (\Delta^{d_0}(y_t - \hat{\alpha}_0))^2} \xrightarrow{w} \\ &\frac{\lambda^2}{\gamma_0} \inf_{\omega \in \Omega} \left(\int_0^1 (B_{d_0}^{\mu_B}(r, \omega))^2 dr \right) \end{aligned}$$

where $B_{d_0}^{\mu_B}(r, \omega)$ is the L_2 projection residual from the continuous time regression,

$$B_{d_0}(r) = \hat{\alpha}_1 + \hat{\delta}_1 dc(\omega, r) + B_{d_0}^{\mu_B}(r, \omega),$$

and $dc(\omega, r) = 1$ if $r > \omega$ and 0 otherwise.

- If $Z_t = \begin{pmatrix} 1 & t \end{pmatrix}$,

$$\begin{aligned} Model 1 : R_b^1 &= T^{1-2d_0} \frac{\inf_{\omega \in \Omega} (\sum (y_t - \hat{\alpha}_1 - (\hat{\alpha}_2 - \hat{\alpha}_1) DC_t - \hat{\beta}_1 t)^2)}{\sum (\Delta^{d_0}(y_t - \hat{\alpha}_0 - \hat{\beta}_0 t))^2} \xrightarrow{w} \\ &\frac{\lambda^2}{\gamma_0} \inf_{\omega \in \Omega} \left(\int_0^1 (B_{d_0}^{\tau_B^1}(r, \omega))^2 dr \right) \end{aligned}$$

$$\begin{aligned} Model 2 : R_b^2 &= T^{1-2d_0} \frac{\inf_{\omega \in \Omega} (\sum (y_t - \hat{\alpha}_1 - \hat{\beta}_1 t - (\hat{\beta}_2 - \hat{\beta}_1) DT_t)^2)}{\sum (\Delta^{d_0}(y_t - \hat{\alpha}_0 - \hat{\beta}_0 t))^2} \xrightarrow{w} \\ &\frac{\lambda^2}{\gamma_0} \inf_{\omega \in \Omega} \left(\int_0^1 (B_{d_0}^{\tau_B^2}(r, \omega))^2 dr \right) \end{aligned}$$

or interest rates.

$$Model \ 3 : \ R_b^3 = T^{1-2d_0} \frac{\inf_{\omega \in \Omega} (\sum (y_t - \hat{\alpha}_1 - (\hat{\alpha}_2 - \hat{\alpha}_1) DC_t - \hat{\beta}_1 t - (\hat{\beta}_2 - \hat{\beta}_1) DT_t)^2)}{\sum \left(\Delta^{d_0}(y_t - \hat{\alpha}_0 - \hat{\beta}_0 t) \right)^2} \xrightarrow{w} \\ \frac{\lambda^2}{\gamma_0} \inf_{\omega \in \Omega} \left(\int_0^1 (B_{d_0}^{\tau_B^3}(r, \omega))^2 dr \right)$$

where $B_{d_0}^{\tau_B^i}$, $i=\{1,2,3\}$ is the L_2 projection residual from the continuous time regressions,

$$\begin{aligned} B_{d_0}(r) &= \hat{\alpha}_1 + \hat{\alpha}_2 dc(\omega, r) + \beta_1 r + B_{d_0}^{\tau_B^1}(r, \omega), \\ B_{d_0}(r) &= \hat{\alpha}_1 + \hat{\beta}_1 r + \hat{\beta}_2 dt(\omega, r) + B_{d_0}^{\tau_B^2}(r, \omega), \\ B_{d_0}(r) &= \hat{\alpha}_1 + \hat{\alpha}_2 du(\omega, r) + \hat{\beta}_1 r + \hat{\beta}_2 dt(\omega, r) + B_{d_0}^{\tau_B^3}(r, \omega), \end{aligned}$$

where $dc(\omega, r)$ is defined as above and $dt(\omega, r) = r - \omega$ for $r > \omega$ and 0 otherwise.

Appendix B gathers the critical values of the distributions above obtained by Monte Carlo simulation for the case where $\Omega = [0.15, 0.85]$ and there is no short-term correlation, that is, $u_t = \varepsilon_t$. When correlation of u_t is suspected, the nuisance parameters λ^2 and γ_0 can be estimated according to the techniques detailed in the previous section.

Notice that if H_0 is rejected, the test favors the hypothesis of $I(0)$ possibly containing breaks. At this stage, the theory for detecting breaks in $I(0)$ processes applies. This theory is more standard and has been studied in depth. For instance, Bai (1997) proposes to use standard t -tests on δ . See also Perron (2005) for a recent survey on this topic.

5. FINITE SAMPLE RESULTS.

To explore the finite sample performance of the test developed in Section 4, the results from some Monte Carlo experiments are reported. In all the experiments, the number of replications was set equal to 5000. Processes were generated according to different DGP's (that will be detailed below). In all of them, innovations were drawn from independent $N(0, 1)$ distributions.

The first experiment was to test the $FI(d)$ hypothesis for several values of $d \in (0.5, 1.5)$, when the true model was generated as the sum of i.i.d innovations plus some deterministic

terms that contained breaks at different time points ($\omega = \{0.20, 0.5, 0.80\}$), according to Models 0 to 3. Different sizes of breaks were considered, both in the constant and in the time trend. In particular, the size of the break in the constant was $\xi_1 = \{0.01, 0.05, 0.1\}$, and in the time trend was $\xi_2 = \{0.005, 0.01, 0.1\}$. No short-term semiparametric correction was introduced to compute the statistics in this case. Very remarkably, the power was equal to 100% in all cases even for moderate sample sizes ($T=100$). Next, short-term correlation was introduced in the DGP. Tables 5.1. to 5-3 present the results of using Models 1-3 to test the hypotheses of interest when the true DGP was an AR(1) process (with an autoregressive coefficient equal to 0.5), plus some breaks. Different locations of the break point were also tried ($\omega = \{0.20, 0.5, 0.80\}$) but only the figures corresponding to $\omega = 0.5$ are reported since they were all very similar.

TABLE 5.1

MODEL 1: POWER R_b^1 TEST; S.L:5%.

True process (H_1) : $y_t = \xi_1 DC_t(\omega) + 0.5y_{t-1} + \varepsilon_t$; $\omega = 0.5$

T=100						T=400				
ξ_1/H_0	$d_0=0.6$	$d_0=0.7$	$d_0=0.8$	$d_0=0.9$	$d_0=1.1$	$d_0=0.6$	$d_0=0.7$	$d_0=0.8$	$d_0=0.9$	$d_0=1.1$
0.01	65.3%	89.5%	98.1%	99.1%	100%	86.3%	93.2%	98.3%	99.2%	100%
0.05	66.9%	89.3%	99.2%	99.1%	100%	86.3%	93.2%	98.6%	100%	100%
0.1	66.6%	90.7%	98.1%	99.0%	100%	86.7%	93.2%	98.2%	100%	100%

TABLE 5.2

MODEL 2: POWER R_b^2 TEST; S.L:5%.

True process (H_1) : $y_t = \xi_2 DT_t(\omega) + 0.5y_{t-1} + \varepsilon_t$; $\omega = 0.5$

T=100						T=400				
ξ_2/H_0	$d_0=0.6$	$d_0=0.7$	$d_0=0.8$	$d_0=0.9$	$d_0=1.1$	$d_0=0.6$	$d_0=0.7$	$d_0=0.8$	$d_0=0.9$	$d_0=1.1$
0.005	61.3%	85.5%	98.1%	99.1%	100%	85.3%	92.7%	97.3%	99.5%	100%
0.01	61.9%	86.3%	99.2%	99.1%	100%	85.2%	92.8%	97.6%	99.3%	100%
0.1	56.6%	84.7%	97.1%	99.0%	100%	81.2%	92.1%	97.3%	99.6%	100%

TABLE 5.3MODEL 3: POWER R_b^3 TEST; S.L:5%.True process $(H_1) : y_t = DC_t(\omega) + \xi_2 DT_t(\omega) + 0.5y_{t-1} + \varepsilon_t; \omega = 0.5$

T=100						T=400				
$(\xi_1, \xi_2) H_0$	$d_0=0.6$	$d_0=0.7$	$d_0=0.8$	$d_0=0.9$	$d_0=1.1$	$d_0=0.6$	$d_0=0.7$	$d_0=0.8$	$d_0=0.9$	$d_0=1.1$
(0.01,0.005)	63.3%	87.5%	97.1%	99.1%	100%	87.3%	92.5%	96.3%	99.6%	100%
(0.05,0.01)	63.9%	87.3%	97.2%	99.1%	100%	87.3%	92.5%	96.3%	99.8%	100%
(0.1,0.1)	56.6%	82.7%	97.1%	99.8%	100%	89.2%	93.0%	95.3%	99.7%	100%

From Tables 5.1 to 5.3 it is seen that the size of the break has not a big impact on the power which, as expected, improves when T and d_0 increase, since the test is consistent and the bigger is d_0 , the more distant H_0 and H_1 are.

Nevertheless, the latter fact should not lead to the wrong rule of always using a large value of d_0 for testing. It should be remembered that if the true DGP is $FI(d^*)$ but the integration order employed to run the test, d_0 , is bigger than the true integration order, d^* , the test would tend to reject the null hypothesis of $FI(d_0)$ even if the true model is fractionally integrated. This fact emphasizes the need of considering fractional hypotheses and not just integer ones, since otherwise it is possible to reject strong persistence when, in fact, it is the source of the non-stationarity in the data. To illustrate the last argument, Table 5.4 presents the rejection frequencies computed from testing $I(1)$ vs. $I(0)$ +breaks when the true DGP is a $FI(d^*)$ model with $d^* = \{0.6, 0.7, 0.8, 0.9\}$ for models 1, 2 and 3. As it can be seen from Table 5.4, rejection of the null hypothesis of $I(1)$ is very likely to occur even for values of d^* close to $d_0 = 1$. This result implies that in applications where only integer values ($I(1)$ vs. $I(0)$ +breaks) are considered, the risk of overrejecting the hypothesis of integration is very high if the true DGP is in fact integrated but of a smaller order than

TABLE 5.4MISSPECIFICATION OF H_0 $H_0 : y_t \sim I(1)$; True DGP: $y_t \sim FI(d^*)$, $d^* < 1$.

	T=100				T=400			
<i>Model</i> / d^*	$d^* = 0.6$	$d^* = 0.7$	$d^* = 0.8$	$d^* = 0.9$	$d^* = 0.6$	$d^* = 0.7$	$d^* = 0.8$	$d^* = 0.9$
Model 1	98.8%	87.2%	57.0%	28.2%	100%	99.8%	91.0%	41.3%
Model 2	92.3%	79.6%	49.6%	24.6%	100%	99.9%	84.2%	34.5%
Model 3	86.4%	60.22%	31.22%	13.25%	100%	99.9%	87.3%	32.5%

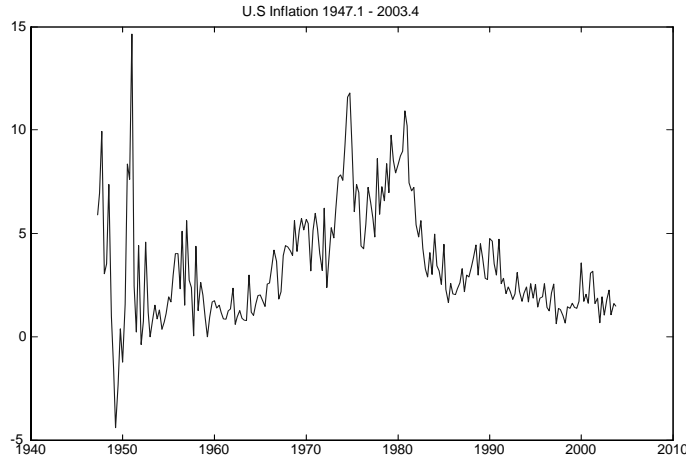
6. EMPIRICAL ILLUSTRATION

The study of the statistical properties of inflation and, in particular, its degree of persistence and stability over time, occupies a privileged place in macro-econometrics since this variable plays a central role in the design of the monetary policy and has important implications for the behavior of private agents. Moreover, a new interest in the subject has arisen in the last few years that has motivated a large number of empirical and theoretical contributions. In spite of this great effort, there is no consensus in the literature about the most appropriate way to model the inflation rate. On the one hand, there is abundant empirical evidence that post-war inflation in industrial countries exhibits high persistence, close to the unit root behavior. The papers of Pivetta and Reis (2004) for the USA and O'Reilly and Whelan (2004) for the euro zone are some examples. On the other, some authors have argued that the above-mentioned results are very sensitive to the employed statistical techniques and that the observed persistence may be due to the existence of unaccounted breaks, probably stemming from changes in the inflation targets of monetary authorities, different exchange rate regimes or shocks in key prices. For instance, Levin and Piger (2003) have found evidence of a break in the intercept of the inflation equation and, conditional on this break, they argue that inflation shows very low persistence. Finally,

¹²Dolado, Gonzalo and Mayoral (2005) provide evidence that this is also the case when other traditional methods (for instance, the Zivot and Andrews (1992) statistic) for testing $I(1)$ vs. $I(0)$ +breaks are employed

Cogley and Sargent (2001, 2003) claim that non-stationary (integrated) representations of inflation are implausible from an economic point of view, since they would imply an infinite asymptotic variance, which could never be optimal if the Central Bank's loss function includes the variance of inflation. Then, they consider inflation as being a short memory ($I(0)$) process.

The aim of this section is to shed further light on this controversy by applying the techniques developed in this article. To facilitate the comparison with previous analysis, the same data set as in Pivetta and Reis (2004) has been employed: The price level, P_t , is measured through the seasonally-adjusted quarterly data on the GDP deflator from the first quarter of 1947 to the last quarter of 2003 (9 observations have been added with respect to their analysis). This data has been obtained from the Bureau of Economic Analysis. Then, inflation is computed as $\pi_t = 400 * \log(P_t/P_{t-1})$, that is, it is the quarterly change of the price level at an annualized rate. Figure 1 presents a plot of this data.



The contradicting results described above could be explained if the inflation rate was a FI process. Unit root tests are known to have very low power against *FI* alternatives. This could account for the non-rejection of this hypothesis in some applications considering inflation. On the other hand, if inflation is FI and standard techniques for detecting and dating breaks are employed, it is well-known that spurious breaks are likely to be detected.

There is both economic and statistical support for the hypothesis of FI in inflation. Gadea and Mayoral (2005) provide an economic justification for the existence of fractional

integration in inflation data. They consider a sticky price model as in Rotemberg (1987), where it is assumed that each firm faces a quadratic cost of changing its price. It is well-known that when this is the case, the dynamics of prices are given by,

$$p_t^i = \vartheta p_{t-1}^i + (1 - \vartheta) p_t^{i*}, \quad (13)$$

where p and p^* represent the actual and optimal level of prices of firm i and the parameter ϑ is a function of the adjustment costs and lies between zero and one. It captures the extent to which imbalances are remedied in each period. Then, if $\vartheta = 1$, there is not price adjustment in period t and if $\vartheta = 0$, the adjustment is perfect. Equation (13) can also be written as:

$$\Delta p_t^i = \vartheta \Delta p_{t-1}^i + \nu_t^i, \quad (14)$$

with $\nu_t^i = (1 - \vartheta) \Delta p_t^{i*}$. Since costs may differ across firms, it is natural to consider the case where ϑ may also depend on i . Then,

$$\Delta p_t^i = \vartheta^i \Delta p_{t-1}^i + \nu_t^i. \quad (15)$$

To build a price index, aggregation over a huge number of individual prices has to be considered.¹³ The change in the price index Δp_t is defined as,

$$\Delta p_t = \sum_{i=1}^N \Delta p_t^i.$$

Application of the aggregation results over heterogeneous agents established in Robinson (1978), Granger (1980), and recently generalized by Zaffaroni (2004) guarantee that if the distribution of ϑ^i verifies some (mild) semi-parametric restrictions, then Δp_t is a $FI(d)$ process.¹⁴ It follows that the higher the proportion of agents correcting the imbalances

¹³For instance, prices for the goods and services used to calculate the CPI are collected in 87 urban areas throughout the United States and from about 23,000 retail and service establishments.

¹⁴Zaffaroni (2004) provides a full discussion of the required restrictions. In particular, ϑ should belong to a family \mathfrak{S} of continuous distributions on $[0,1)$ with density,

$$\mathfrak{S}(\vartheta, d) \sim c\vartheta^{-d} \text{ as } \vartheta \rightarrow 0^+ \quad (16)$$

between the actual and the optimal level of prices by a small amount each (i.e., $\vartheta^i \approx 1$), the higher the inflation inertia.

From an applied point of view, evidence in favor of FI behaviour in inflation data has been reported in several papers (see, among others, Delgado and Robinson (1994), Baillie, Chung and Tieslau (1992, 1996), Backus and Zin, (1993), etc.). Nevertheless, the methods employed in these contributions are not robust to the existence of structural breaks. Therefore, it remains to check whether the evidence supporting FI can be due to the existence of structural breaks.

To begin the analysis, Table 6.1 presents the results of some standard tests for unit roots. The first two columns contain the figures from the Augmented Dickey-Fuller (ADF) and the Phillips- Perron (P-P) tests of $I(1)$ vs. $I(0)$ and the third one, those obtained by applying the KPSS test of $I(0)$ vs. $I(1)$. From the first two columns it is seen that when the unit root model is tested against $I(0)$, the former is rejected. The opposite result is obtained when the hypothesis are reversed (third column). In this case, also the $I(0)$ is rejected against the alternative of $I(1)$.

TABLE 6.1

UNIT ROOT TESTS

	ADF	P-P	KPSS ($I(0)$ vs. $I(1)$)
Value of the test	-3.49**	-5.88**	0.67*
Critical Values (5%)	-2.87		0.463

*, **Rejection at the 5% and the 1% level, respectively.

The rejection of the $I(1)$ and the $I(0)$ hypotheses is compatible with the existence of both fractional integration and also with some types of structural breaks. This is so because unit

with $c \in (0, \infty)$. Many parametric specifications verify the above-mentioned restrictions, such as the uniform or the Beta distribution. It is interesting to notice that the behavior of $\mathfrak{S}(\vartheta, d)$ within any interval $[0, \gamma]$ is completely unspecified.

root tests are known to have some power against the latter DGP's (see Lee and Schmidt (1996), Diebold and Rudebush (1991) and Perron (1989)).

The next step is to test for the suitability of the FI specification. Table 6.2 presents the results of estimating d using different techniques: the semiparametric estimator proposed by Geweke and Porter Hudak (GPH , (1982)), Non-linear least squares (NLS , see Beran (1995)), Exact Maximum likelihood (EML, Sowell (1992)) and Minimum Distance (MD, Mayoral (2004)). In all cases, fractional values of d 'far' from both the $I(0)$ and the $I(1)$ hypothesis are found.

TABLE 6.2

ESTIMATION OF d				
	GPH	NLS	EML	MD
\hat{d}	0.51** (0.13)	0.59** (0.15)	0.61** (0.14)	0.58** (0.12)

Tests of fractional versus integer integration ($d = 0$ or 1) based on the values above are not able to reject the FI hypothesis at the 5% signification level, confirming previous findings about the existence of FI in inflation data (see the papers cited above). Nevertheless, it has also been argued that estimates of d can be very imprecise¹⁵ and that tests of integer versus fractional integration based on these estimates in general do not posses good properties. Table 6.3 reports the results of testing the null hypothesis of $I(1)$ versus the hypothesis of fractional integration using both Wald-type techniques (the augmented Fractional Dickey-Fuller test, see Dolado, Gonzalo and Mayoral, (2002,2004)) and LM ones (Tanaka, 1999). In both cases the unit root hypothesis is rejected in favor of fractional integration.

¹⁵Parametric methods are very sensitive to the specified models and semiparametric ones are known to be biased in the presence of strong short-term autocorrelation.

TABLE 6.3TESTS OF $I(1)$ VS $FI(d)$.

d – Number of lags	AUG. FDF ¹⁶				LM TEST
$H_1 :$	d=0.6	d=0.7	d=0.8	d=0.9	$d < 1$
$d_0 = 0.7$	-3.87**	-3.76**	-3.61**	-3.46**	-2.85**
Critical Values (5%)			-1.65		

Finally, it remains to be checked whether the evidence in favor of FI can be due to the existence of breaks in the intercept of the inflation equation as claimed by Levin and Piger, (2003). Table 6.4 presents the values of testing FI against a short-memory process that may contain a break in the intercept. Different values of d and also for different values of lags to compute the Newey-West correction have been considered.

TABLE 6.4LR TESTS $FI(d)$ VS $I(0)$ WITH ONE BREAK IN THE CONSTANT.

d – Number of lags*	1	2	3	Crit. Values S.L.:5%
$d_0 = 0.6$	0.840	1.282	1.492	0.3958
$d_0 = 0.7$	0.294	0.432	0.497	0.1752
$d_0 = 0.8$	0.102	0.162	0.182	0.0841
$d_0 = 0.9$	0.035*	0.043	0.052	0.0403
$d_0 = 1$	0.011*	0.012*	0.014*	0.0200

The suggested number of lags, following Andrews (1991) method was 3. These implies that the null hypothesis of FI cannot be rejected against the alternative of $I(0)+$ breaks. It is remarkable, however, that the null hypothesis of $I(1)$ (permanent shocks) is rejected.

Summarizing, we have found evidence supporting the hypothesis of FI in inflation data and we have checked that this specification is preferred to another one containing a break in the intercept of the inflation equation. This finding has important implications when

¹⁶The number of lags to compute the AFDF test and to select the parametric model for the LM test have been chosen according to the BIC criterion.

computing estimates of the persistence of the process. Impulse responses of inflation computed on $I(1)$ specifications will deliver biased estimates that, specially in the medium and long term, will tend to overestimate persistence if the process is $FI(d)$, with $d < 1$. Moreover, other popular tools for analyzing persistence such as the sum of the autoregressive coefficients could easily lead to wrong conclusions. Under this approach, if the sum of AR coefficients is close to 1, the process is said to contain an (integer) unit root and therefore is considered to be very persistent. But once FI is allowed for, this conclusion is clearly wrong since the sum of the autoregressive coefficients of a $FI(d)$ process is equal to one for any $d > 0$. See Gadea and Mayoral (2005) for a more detailed discussion of this issue.

7. CONCLUSIONS

This paper analyzes the long-standing issue of determining the source of the non-stationarity observed in many economic variables: whether it is a result of a high degree of inertia (very persistent shocks) or it appears as a consequence of the existence of rare and unexpected events that are able to change the underlying structure of the series (breaks). We have extended the traditional approach of testing $I(1)$ versus $I(0)$ +breaks by allowing for a richer class of persistent behaviors under H_0 . In particular, the possibility of fractional integration has been explicitly taken into account. It has also been shown that explicitly considering FI processes is very relevant since tests of $I(1)$ vs. $I(0)$ +breaks tend to reject the former hypothesis when the true DGP is a FI process with an integration order smaller than 1. The asymptotic properties of the tests statistics as well as their finite sample behavior have been analyzed. Finally, an empirical application that analyzes US inflation has been reported and evidence of FI behavior has been found in this data set. This finding helps to reconcile previous controversies that exist in the literature.

REFERENCES

- Akonom J. and C. Gouriou (1987), "A functional central limit theorem for fractional processes", Discussion Paper 8801, CEPREMAP, Paris.
- Andrews, D. W. K. (1991), "Heteroskedasticity and autocorrelation consistent covariance matrix estimation", *Econometrica* 59, 817-854.
- Andrews, D. W. K. (1993), "Test for parameter instability and structural change with unknown break point", *Econometrica* 61, N ° 4, 821-856.
- Bai, J. (1997), "Estimation of a change point in multiple regression models", *The Review of Economics and Statistics*, Vol. 79, N ° 4, 551-563.
- Bai, J. (1999), "Likelihood ratio tests for multiple structural changes", *Journal of Econometrics*, Vol. 91, N ° 2, 299-323.
- Bai J. and P. Perron (1998), "Testing for and estimation of multiple structural changes", *Econometrica* 66, 47-79.
- Bai J. and P. Perron (2003), "Critical values of Multiple Structural Change Tests", *Econometrics Journal*, 6, 72-78.
- Baillie, R.T. (1996): "Long memory processes and fractional integration in economics and finance," *Journal of Econometrics*, 73, 15-131.
- Baillie R., Ch. F. Chung and M. Tieslau (1992), "The long memory and variability of inflation: a reappraisal of the Friedman hypothesis", Discussion paper, Tilburg University.
- Baillie R., Ch. F. Chung and M. Tieslau (1996), "Analysing inflation by the fractionally integrated ARFIMA-GARCH model", *Journal of Applied Econometrics*, 11, 23-40.
- Backus, D. K, and S. E. Zin, (1993), "Long memory inflation uncertainty: evidence from the term structure of interest rates", *Journal of Money, Credit and Banking*, 25, 681-700.
- Bhargava, A. (1986), "On the theory of testing for unit roots in observed time series", *The Review of Economic Studies* 53 (3), 369-384.
- Bhattacharya, R.N., Gupta, V.K., E. Waymire (1983), "The Hurst effect under trends". *Journal of Applied Probability* 20, 649-662.

Beran, J. (1995), "Maximum likelihood estimation of the differencing parameter for invertible and short and long memory autoregressive integrated moving average models". *Journal of the Royal Statistical Society, Series B*, 57, No. 4, 659-672.

Cogley, T. and T. J. Sargent (2004), "The conquest of US inflation. Learning and robustness to model uncertainty". Mimeo.

Cogley, T. and T. J. Sargent (2001), "Evolving post World War II inflation dynamics". NBER Macroeconomics Annual 16, 331-373.

Davidson, J. and P. Sibbertsen (2003), "Generating schemes for long memory processes: regimes, aggregation and linearity", Mimeo.

Delgado and Robinson (1994), "New Methods for the Analysis of Long-memory Time Series", *Journal of Forecasting* 13, 97-107.

Diebold, F.X. and A. Inoue (2001), "Long memory and regime switching", *Journal of Econometrics*, 105, 131-159.

Diebold, F.X. and G. Rudebush (1991), "On the power of the Dickey-Fuller tests against fractional alternatives", *Economic Letters*, 35, 55-160.

Ding, Z., Granger, C. W. J. and R.F. Engle (1993). "A long memory property of stock market returns and a new model", *Journal of Empirical Finance*, vol. 1(1), 83-106.

Dolado J.J., J. Gonzalo and L. Mayoral (2002), "A Fractional Dickey-Fuller test for unit roots", *Econometrica* 70, 5.

Dolado J.J., J. Gonzalo and L. Mayoral (2003), "Long range dependence in Spanish political opinion poll data", *Journal of Applied Econometrics*, Vol. 18, N^o 2, p. 137-155.

Dolado J.J., J. Gonzalo and L. Mayoral (2004), "Testing for a unit root against fractional alternatives in the presence of a maintained trend", Mimeo.

Dolado J.J., J. Gonzalo and L. Mayoral (2005), "What is what? a simple time-domain test of long-memory vs. structural breaks", In preparation.

Dufour, J. and M.L. King (1991), "Optimal invariant tests for the autocorrelation coefficient in linear regressions with stationary and nonstationary AR(1) errors", *Journal of Econometrics* 47, 115-143.

Elliott G., E. Rothemberg and J. Stock (1996), "Efficient test for an autoregressive unit

roots", *Econometrica* Vol. 64, N^o 4.

Gadea, L. and L. Mayoral (2005), "The persistence of inflation in industrial countries". Mimeo.

Geweke, J., and S. Porter-Hudak (1983), "The estimation and application of long memory time series models," *Journal of Time Series Analysis*, 4, 221-238.

Giraitis L., P. Kokoszka and R. Leipus (2001), "Testing for long memory in the presence of a general trend". *Journal of Applied Probability*, 38, No 4, 1033-1054.

Granger C.W.J (1966), "The typical spectral shape of an economic variable", *Econometrica* 34, 50-161.

Granger C.W.J. and N. Hyung (1999), "Occasional structural breaks and long memory", Working Paper, University of California, San Diego.

Granger, C.W.J and K. Joyeux (1980), "An introduction to long memory series", *Journal of Time series Analysis*, 1, 15-30.

Granger, C.W.J. (1980), "Long memory relationships and the aggregation of dynamic models", *Journal of Econometrics* 14, 227-238.

Grenander U. and M. Rosemblatt (1984), *Statistical Analysis of Stationary Time Series*, 2nd edition. Chelsea Publishing Co., New York.

Hannan, E.J. (1976), "The asymptotic distribution of serial covariances", *Annals of Statistics* 4, 396-399.

Hansen, B. (2001), "The new econometrics of structural change: Dating the changes in US labor productivity", *Journal of Economic Perspectives*, 15, 117-128.

Henry, M. and P. Zaffaroni (2002), "The long range dependence paradigm for Macroeconomics and Finance", in *Long range dependence: Theory and applications*, P. Doukhan, G. Oppenheim and M. Taqqu (ed). Birkhäuser, Boston.

Hidalgo, J. and P. Robinson, (1996), "Testing for Structural change in a long-memory environment", *Journal of Econometrics*, 70, (1), 159-174.

Hidalgo, J. and P. Robinson, (1997), "Time Series Regression with Long Range Dependence", *Annals of Statistics*, 25, 2054-2083.

Hosking, J. R.M. (1981), "Fractional differencing", *Biometrika* 68, 165-176.

- Hosking, J. R.M., (1996), "Asymptotic distributions of the sample mean, autocovariances and autocorrelations of long-memory time series", *Journal of Econometrics* 73, 261-284.
- Hsu, C. C. (2001), "Change point estimation in regressions with $I(d)$ variables", *Economics Letters* 70, 147-155.
- Kramer W. and P. Sibbertsen (2002), "Testing for structural changes in the presence of long memory", *International Journal of Business and Economics*, Vol.1, N^o 3, 235-242.
- Kwiatkowski, D., P.C.B. Phillips, P. Schmidt and Y. Shin, "Testing the null hypothesis of stationarity against the alternative of a unit root", *Journal of Econometrics*, 54, 159-178.
- Künsh, H. (1986), "Discrimination between deterministic trends and long range dependence", *Journal of Applied Probability* 23, 1025-1030.
- Lazarova, S. (2004), "Testing for structural change in regression with long memory processes", forthcoming in *The Journal of Econometrics*.
- Lehmann E. L. (1959), *Testing Statistical Hypothesis*. John Wiley and Sons: New York.
- Levin A. and J. Piger (2003), "Is inflation persistence intrinsic in industrial economics", Federal Reserve Bank of Saint Louis, Working Paper 2002-023.
- Lee, D and P. Schmidt, (1996), "On the power of the KPSS test of stationarity against fractionally-integrated alternatives", *Journal of Econometrics* 73, 285-302.
- Lobato, I.N., and N. E. Savin, (1998), "Real and spurious long-memory properties of stock-market data", *Journal of Business and Economic Statistics*, Vol. 16, N^o 3, 261-283.
- Lucas, R.E (1976), "Econometric policy evaluation: A critique", *Carnegie-Rochester conference on public policy* 1, 19-46.
- Marinucci, D. and P. Robinson, (1999), "Alternative forms of fractional Brownian motion", *Journal of Statistical Planning and Control*, 80, p. 111-122.
- Marmol, F. and C. Velasco (2002), "Trend stationarity versus long-range dependence in time series analysis", *Journal of Econometrics* 108, 25-42.
- Mayoral, L. (2004), "A new minimum distance estimator for ARFIMA processes", Working paper, Universitat Pompeu Fabra.
- Mikosch, T. and Starica, C. (2000), "Limit theory for the sample autocorrelations and extremes of a GARCH(1,1) process", *Annals of Statistics* 28, 1427-1451.

- Mikosch, T. and Starica, C. (2004), “Non-stationarities in financial time series, the long range dependence and the IGARCH effects”, *Review of Economics and Statistics*, 86(1), 378-390.
- Ming, L. (1998), “Asymptotics of nonstationary fractional integrated series”, *Econometric Theory* 14, 641-662.
- Nelson C. R. and C.I. Plosser, (1982), “Trends and Random walks in macroeconomic time series: Some evidence and implications”, *Journal of Monetary Economics* 10, 139-162.
- Nunes, L. C., C. M. Kuan and P. Newbold (1995), “Spurious Breaks”, *Econometric Theory* 11, 736-749.
- Perron, P. (1989), “The great crash, the oil price shock and the unit root hypothesis”, *Econometrica* 58, 1361-1401.
- Perron P. and Z. Qu (2004), “An analytical evaluation of the log-periodogram estimate in the presence of level shifts and its implications for stock returns volatility”. Mimeo.
- Phillips P.C.B. and Z. Xiao, (1998), “A primer on unit root testing”, *Journal of Economic Surveys*, Vol. 12, N^o 5, p.423-469.
- Pivetta F. and R. Reis, (2004), “The persistence of inflation in the United States”, Mimeo.
- Perron, P., (2005), “Dealing with Structural Breaks”, forthcoming in the *Palgrave Handbook of Econometrics*, Volume 1: Econometric Theory
- Robinson, P.M. (1994), “Efficient tests on nonstationarity hypotheses”. *Journal of the American Economic Association* 89, 1420-1437.
- Robinson, P. (2004), *Time series with Long Memory*, Oxford University Press.
- Rotemberg, J. (1987), “The New Keynesian Microfoundations”, *Macroeconomics Annual*, 2, 69-104.
- Sargan J.D. and A. Bhargava (1983), “Testing residuals from least squares regression for being generated by the gaussian random walk”, *Econometrica* 51 (1), 153-174.
- Sibbertsen P. and I. Venetis (2004), “Distinguishing between long-range dependence and deterministic trends”, Mimeo.
- Sowell, F.B.(1992): “Maximum likelihood estimation of stationary univariate fractionally-integrated time-series models,” *Journal of Econometrics*, 53, 165-188.

Tanaka, K. (1999), "The nonstationary fractional unit root," *Econometric Theory*, 15, 549-582.

Teverosky V. and M. Taqqu (1997), "Testing for long-range dependence in the presence of shifting means or a slowly declining trend, using a variance-type estimator", *Journal of Time Series Analysis*, Vol. 18, N^o 3, 279-304.

Zivot, E. and D.W.K. Andrews (1992), "Further evidence on the Great Crash, the oil-price shock and the unit root hypothesis", *Journal of Business and Economic Statistics* Vol. 10, N^o 3, 251-270.

APPENDIX A

Proof of Theorem 1

1. Under the null hypothesis, the process $\Delta^{d_0} y_t$ equals ε_t and then a standard LLN for *i.i.d* process guarantees that $\sum (\Delta^{d_0} y_t)^2 / T \xrightarrow{p} \sigma^2$. On the other hand, since under the null $\Delta^{d_1} y_t \sim FI(d_0 - d_1)$, with $(d_0 - d_1) > 0.5$, then $T^{-2(d_0 - d_1)} \sum \Delta^{d_1} y_t^2 \xrightarrow{w} \sigma^2 \int_0^1 B_{(d_0 - d_1)}^2(r) dr$, (see Akonon and Gouriéroux, (1987)).
2. Under the null hypothesis, the variance of the (stationary) process y_t is given by $\sigma^2 \frac{\Gamma(1-2(d_0 - d_1))}{\Gamma^2(1-(d_0 - d_1))}$ (see Baillie, (1996)). Hosking (1996) provides for the asymptotic distribution of $\sum \left(\Delta^{d_1} y_t^2 - \sigma^2 \frac{\Gamma(1-2(d_0 - d_1))}{\Gamma^2(1-(d_0 - d_1))} \right)$. Whenever $(d_0 - d_1) > 1/4$, it is given by the Roseblatt distribution with cumulants defined in equations (5) and (6) in Hosking (1996). If $(d_0 - d_1) < 1/4$, the distribution is normal and the variance is defined in Hannan (1976). ■

Proof of Theorem 2

The asymptotic behaviour of the numerator depends upon the distance between the true integration order of the process, d^* , and d_1 , the order employed to run the test under the alternative hypothesis. Three different situations have to be distinguished.

1. $(d^* - d_1) > 0.5$.

Following the same reasoning as in the proof of Theorem 1, the denominator is $O_p(T)$, since $\Delta^{d_0} y_t \sim FI(d^* - d_0)$ is stationary and verifies a LLN. On the other hand the process $\Delta^{d_1} y_t$ is a (nonstationary) $FI(d^* - d_1)$, and therefore $\sum \Delta^{d_1} y_t^2 = O_p(T^{2(d^* - d_1)})$. Then, $T^{1-2(d_0 - d_1)} \sum \Delta^{d_1} y_t^2 / \sum \Delta^{d_0} y_t^2$ is $O_p(T^{-2(d_0 - d^*)})$, which tends to zero as long as $d_0 > d^*$.

2. $(d^* - d_1) = 0.5$.

Notice first that, since $d_0 > d^*$, then $(d_0 - d_1) > 0.5$. Also, since $\Delta^{d_1} y_t \sim FI(0.5)$, it follows that $\sum \Delta^{d_1} y_t^2 = O_p(T \log T)$, (see Ming (1998)). Then, $T^{1-2(d_0 - d_1)} \sum \Delta^{d_1} y_t^2 / \sum \Delta^{d_0} y_t^2$ is $O_p(T^{1-2(d_0 - d_1)}(\log(T)))$, which tends to zero as long as $d_0 > d_1$.

3. $(d^* - d_1) < 0.5$.

Both the numerator and the denominator are (asymptotically) stationary and ergodic now, therefore $\sum \Delta^{d_0} y_t^2 / \sum \Delta^{d_1} y_t^2 = O_p(1)$. Notice that in this case $(d_0 - d_1) \leq 0.5$. If d_0 is chosen such that $(d_0 - d_1) > 0.5$, it turns out that the test tends to zero at a rate $T^{1-2(d_0-d_1)}$. On the other hand, if d_0 is chosen such that $(d_0 - d_1 < 0.5)$, it is easy to check that $\lim_{T \rightarrow \infty} \sum \Delta^{d_1} y_t^2 / \sum \Delta^{d_0} y_t^2 < \frac{\Gamma(1-2(d_0-d_1))}{\Gamma^2(1-(d_0-d_1))}$, therefore $T^A \left(\sum \Delta^{d_1} y_t^2 / \sum \Delta^{d_0} y_t^2 - \frac{\Gamma(1-2(d_0-d_1))}{\Gamma^2(1-(d_0-d_1))} \right)$, (where $A = 1/2$ or $1 - 2(d_0 - d_1)$ according to whether $(d_0 - d_1)$ is greater or smaller than $1/4$) diverges to $-\infty$, so that the null hypothesis will be rejected with probability 1. ■

Proof of Theorem 3

The proof of this theorem is straight forward given Theorem 1 and the results in Akonom and Gouriéroux (1987). ■

Proof of Theorem 4

1. The denominator of R^c divided by T converges to γ_0 since $\hat{\alpha}_0$ is consistently estimated.

With respect to the numerator, notice that $\hat{\alpha}_1 = \sum_{t=1}^T y_t / T$, and then,

$$\sum_{t=1}^T (y_t - \hat{\alpha}_1)^2 = \sum_{t=1}^T \left(\alpha_1 + x_t - \alpha_1 - T^{-1} \sum_{t=1}^T x_t \right)^2 \quad (17)$$

$$= \sum_{t=1}^T x_t^2 - T^{-1} \left(\sum_{t=1}^T x_t \right)^2, \quad (18)$$

where x_t is a $\text{FI}(d_0)$ process as in (3) that does not contain any deterministic component. Application of the continuous mapping theorem delivers,

$$\begin{aligned} T^{-2d_0} \sum_{t=2}^T (y_t - \hat{\alpha}_1)^2 &= T^{-2d_0} \sum_{t=1}^T x_t^2 - T^{-1} \left(\sum_{t=1}^T x_t \right)^2 \\ &\xrightarrow{w} \lambda^2 \left(\int_0^1 (B_{d_0}^\mu(r))^2 dr \right) \end{aligned}$$

where $B_{d_0}^\mu(r) = B_{d_0}(r) dr - \left(\int_0^1 B_{d_0}(r) dr \right)$ is a *demeaned fBM*.

2. Since the test is invariant to the true values of the deterministic components, let us assume without loss of generality that the true *DGP* does not contain these

componentes. From Marmol and Velasco (2002) it is known that,

$$T^{-2d_0} \sum \left(y_t - \hat{\alpha}_1 - \hat{\beta}_1 t \right)^2 \xrightarrow{w} \lambda^2 \left(\int_0^1 B_{d_0}^2(r) dr - \left(\int_0^1 B_{d_0}(r) dr \right)^2 - 12 \left(\int_0^1 (r - 1/2) B_{d_0}(r) dr \right)^2 \right) \quad (19)$$

As in the unit root case, the detrended fBM is given by $B_{d_0}^T(r) = B_{d_0}(r) dr - 4 \left(\int_0^1 B_{d_0}(s) ds - 3/2 \int_0^1 s B_{d_0}(s) ds \right) + 6r \left(\int_0^1 s B(s) ds - 2 \int_0^1 s B(s) ds \right)$.

On the other hand, it is straight forward to show that under the null hypothesis:

$$T^{-1} \sum \left(\Delta^{d_0} y_t - \hat{\alpha}_0 - \hat{\beta}_0 t \right)^2 \xrightarrow{p} \gamma_0. \quad (20)$$

Combining equations (19) and (20) the desired result is obtained. ■

Proof of Theorem 5

The proof of this theorem can be constructed along the lines of that of Theorem 1 in Zivot and Andrews (1992) (Z&A henceforth). As they point out, there exist several ways of proving this type of results. One way is to prove the weak convergence of the proposed test statistics to some process $L(\cdot)$ and then, provided $\inf_{\omega \in \Omega} L(\omega)$ is a continuous functional of $L(\cdot)$, to apply the continuous mapping theorem (CMT) to obtain the desired result. But, in order to avoid the difficulty of establishing tightness (which is required in order to show weak convergence), another method of proof will be used.

Following the notation in Z&A, let us define $z_{tT}^i(\omega)$ for $i = \{0, 1, 2, 3\}$ as the vector that contains the deterministic components for each model under the alternative hypothesis. For instance if $i = 1$, $z_{tT}^1(\omega)' = \begin{pmatrix} 1 & t & DC_t(\omega) \end{pmatrix}$. We will also need a rescaled version of the deterministic regressors, $\tilde{z}_T^i(\omega, r) = \delta_T^i z_{[Tr]T}^i(\omega)$, where δ_T^i is a diagonal matrix of weights¹⁷. The test statistics can be rewritten as:

¹⁷For instance, in Model 1,

$$\delta_T^1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & T & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

$$\inf_{\omega \in \Omega} R_b^i(\omega) = \inf_{\omega \in \Omega} \frac{\sum_{i=1}^T (y_t^i(\omega))^2}{\sum_{i=2}^T (\Delta^{d_0} y_t^i)^2}, \text{ for } i = \{0, 1, 2, 3\}, \quad (21)$$

where $y_t^i = y_t - z_{tT}^i(\omega)' \left(\sum_{s=1}^T z_{sT}^i(\omega) z_{sT}^i(\omega)' \right)^{-1} \sum_{s=1}^T z_{sT}^i(\omega) y_s$ for $i = \{0, 1, 2, 3\}$, $\Delta^{d_0} y_t^0 = \Delta^{d_0} y_t - \hat{\alpha} \Delta^{d_0}$ and $\Delta^{d_0} y_t^i = \Delta^{d_0} y_t - \hat{\alpha} \Delta^{d_0} - \hat{\beta} \Delta^{d_0-1}$ for $i = \{1, 2, 3\}$. Henceforth, only Model 1 will be considered. Proofs for models $\{0, 2, 3\}$ are analogous and therefore, are omitted.

It is straight forward to check that under H_0 , $T^{-1} \sum_{i=2}^T (\Delta^{d_0} y_t^1)^2 \xrightarrow{p} \sigma^2$. Now we will derive the limiting distribution of the numerator of (21). The proof will be completed in three steps, that will closely follow Z&A's approach. In the first one, it will be shown the numerator in (21) can be written as a functional $g(.,.)$ of the partial sum process $X_T(.)$ defined as,

$$X_T(r) = T^{1/2-d} \sigma^{-1} \sum_{i=0}^{[Tr]} \pi_i(-d) \varepsilon_{[Tr]-i}, \quad (j-1) < r < (j+1) \text{ for } j = 1, \dots, T,$$

and a rescaled version of the deterministic components, $\tilde{z}_T(.,.)$. In the second step, it is needed to show that $(X_T(.), \tilde{z}_T(.,.))$ jointly converge to $(B_d(.), \tilde{z}(.,.))$. Finally, it will be checked that g is continuous with respect to $(B_d(.), \tilde{z}(.,.))$ and then, convergence of the statistics would follow by applying the CMT (third step).

First step. By expression (A.5) in Z&A,

$$\begin{aligned} T^{-2d} \inf_{\omega \in \Omega} \sum_{i=1}^T (y_t^i(\omega))^2 &= \int_0^1 \left\{ \sigma X_T(r) - \tilde{z}_T(\omega, r)' \left(\int_0^1 \tilde{z}_T(\omega, s)' \tilde{z}_T(\omega, s)' ds \right)^{-1} \right. \\ &\quad \times \left. \left(\int_0^1 \tilde{z}_T(\omega, s)' \sigma X_T(s)' ds \right) \right\}^2 dr + o_{p\omega}(1) \\ &= g[\sigma X_T, \tilde{z}_T](\omega) + o_{p\omega}(1). \end{aligned}$$

Second step. By Akonom and Gourioux (1987),

$$T^{-2d_0} X_T(.) \xrightarrow{w} B_{d_0}(.)$$

and by A&Z,

$$\tilde{z}_T^1(.,.) \rightarrow z^1(.,.)^{18}.$$

¹⁸For instance, $\tilde{z}_T^1(\omega, r) \rightarrow \tilde{z}^1(\omega, r) = \begin{pmatrix} 1 & r & du(\omega, r) \end{pmatrix}'$.

Since the limiting distribution of $\tilde{z}_T(.,.)$ is degenerate, it follows that $(X_T(.), \tilde{z}_T(.,.))$ converge weakly to $(B_d(.), z(.,.))$ ¹⁹.

Third step. The final step is to show that g defines a continuous functional with probability 1 with respect to the limit process $(B_d(.), z(.,.))$. This is done in Z&A through a series of steps (Lemmas A.2-A.4). The continuity of g follows from the continuity of a composition of continuous functionals and the result of the theorem follows from the CMT. ■

¹⁹The uniform metric is used in the first term whereas the d^* metric is used in the second. See Z&A for details.

APPENDIX B

TABLE B1

Critical values R test

Simplest case: $H_0 : \Delta^{d_0} y_t = \varepsilon_t$; $H_1 : \Delta^{d_1} y_t = \varepsilon_t$

$(d_0 - d_1) T$	$T = 100$			$T = 400$			$T = 1000$		
$(d_0 - d_1)$	10%	5%	1%	10%	5%	1%	10%	5%	1%
0.6	0.6681	0.6014	0.5066	0.7552	0.6832	0.5733	0.8156	0.7381	0.6290
0.7	0.3551	0.3043	0.2359	0.3752	0.3231	0.2564	0.3848	0.3315	0.2649
0.8	0.2032	0.1665	0.1231	0.2043	0.1670	0.1184	0.2010	0.1644	0.1180
0.9	0.1248	0.0962	0.0650	0.1212	0.0950	0.0637	0.1209	0.0944	0.0632
1.0	0.0793	0.0590	0.0378	0.0769	0.0566	0.0352	0.0752	0.0512	0.0342
1.1	0.0543	0.0382	0.0215	0.0510	0.0358	0.0192	0.0511	0.0359	0.0205
1.2	0.0376	0.0257	0.0133	0.0357	0.0244	0.0125	0.0364	0.0246	0.0125
1.3	0.0267	0.0178	0.0088	0.0254	0.0161	0.0082	0.0242	0.0160	0.0082
1.4	0.0197	0.0122	0.0058	0.0184	0.0117	0.0051	0.0185	0.0115	0.0051

TABLE B2
Critical values R^c test

Simplest case: $H_0 : \Delta^{d_0}(y_t - \alpha_0) = \varepsilon_t$; $H_1 : (y_t - \alpha_1) = \varepsilon_t$

$d_0 T$	$T = 100$			$T = 400$			$T = 1000$		
d_0	10%	5%	1%	10%	5%	1%	10%	5%	1%
0.6	0.5471	0.5074	0.4371	0.6366	0.5904	0.5159	0.6888	0.6426	0.5615
0.7	0.2701	0.2400	0.2003	0.2854	0.2563	0.2150	0.2958	0.2656	0.2203
0.8	0.1416	0.1228	0.0964	0.1434	0.1232	0.0944	0.1446	0.1248	0.0961
0.9	0.0797	0.0657	0.0487	0.0778	0.0640	0.0456	0.0775	0.0645	0.0470
1.0	0.0474	0.0381	0.0260	0.0452	0.0331	0.0242	0.0444	0.0300	0.0221
1.1	0.0298	0.0230	0.0148	0.0288	0.0215	0.0135	0.0291	0.0224	0.0148
1.2	0.0204	0.0151	0.0089	0.0193	0.0142	0.0083	0.0191	0.0141	0.0078
1.3	0.0138	0.0098	0.0051	0.0130	0.0093	0.0049	0.0127	0.0091	0.0051
1.4	0.0098	0.0066	0.0034	0.0093	0.0061	0.0031	0.0094	0.0063	0.0032

TABLE B3**Critical values R^τ test**Trended case: $H_0 : \Delta^{d_0}(y_t - \alpha_0 - \beta_0 t) = \varepsilon_t$; $H_1 : y_t - \alpha_1 - \beta_1 t = \varepsilon_t$

$d_0 T$	$T = 100$			$T = 400$			$T = 1000$		
d_0	10%	5%	1%	10%	5%	1%	10%	5%	1%
0.6	0.4864	0.4513	0.3948	0.5740	0.5351	0.4770	0.6258	0.5875	0.5255
0.7	0.2243	0.2053	0.1755	0.2417	0.2214	0.1925	0.2512	0.2304	0.1963
0.8	0.1100	0.0978	0.0815	0.1107	0.0984	0.0798	0.1122	0.0997	0.0830
0.9	0.0550	0.0478	0.0373	0.0540	0.0470	0.0369	0.0541	0.0470	0.0367
1.0	0.0293	0.0248	0.0179	0.0283	0.0239	0.0173	0.0277	0.0230	0.0169
1.1	0.0159	0.0132	0.0095	0.0154	0.0127	0.0089	0.0155	0.0126	0.0090
1.2	0.0093	0.0075	0.0051	0.0089	0.0070	0.0049	0.0087	0.0068	0.0046
1.3	0.0054	0.0042	0.0028	0.0051	0.0040	0.0024	0.0051	0.0039	0.0025
1.4	0.0033	0.0024	0.0015	0.0031	0.0024	0.0015	0.0030	0.0022	0.0014

TABLE B4

Critical values R_b^0 test

Model 0: $H_0 : \Delta^{d_0}(y_t - \alpha_0) = \varepsilon_t$; $H_1 : y_t - \alpha_1 - \alpha_2 DC_t(\omega) = \varepsilon_t$

$d_0 T$	$T = 100$			$T = 400$			$T = 1000$		
d_0	10%	5%	1%	10%	5%	1%	10%	5%	1%
0.6	0.4231	0.3992	0.3503	0.5194	0.4862	0.4325	0.5756	0.5403	0.4995
0.7	0.1891	0.1753	0.1584	0.2135	0.1973	0.1765	0.2246	0.2097	0.1852
0.8	0.0941	0.0844	0.0692	0.0966	0.0880	0.0733	0.0960	0.0874	0.0733
0.9	0.0452	0.0404	0.0335	0.0454	0.0309	0.0381	0.0463	0.0407	0.0293
1.0	0.0244	0.0212	0.0159	0.0238	0.0200	0.0158	0.0234	0.0190	0.0156
1.1	0.0141	0.0124	0.0081	0.0131	0.0111	0.0082	0.0132	0.0112	0.0071
1.2	0.0083	0.0079	0.0052	0.0083	0.0065	0.0042	0.0082	0.0064	0.0042
1.3	0.0053	0.0043	0.0034	0.0032	0.0043	0.0054	0.0051	0.0043	0.0023
1.4	0.0037	0.0025	0.0011	0.0024	0.0024	0.0034	0.0033	0.0023	0.0011

TABLE B5
Critical values R_b^1 test

Model 1: $H_0 : \Delta^{d_0}(y_t - \alpha_0 - \beta_0 t) = \varepsilon_t$; $H_1 : y_t - \alpha_1 - \alpha_2 DC_t(\omega) - \beta_1 t = \varepsilon_t$

$d_0 T$	T=100			T=400			T=1000		
d_0	10%	5%	1%	10%	5%	1%	10%	5%	1%
0.6	0.3901	0.3722	0.3393	0.4816	0.4570	0.4174	0.5348	0.5110	0.4716
0.7	0.1712	0.1601	0.142	0.1907	0.1791	0.1592	0.2001	0.1867	0.1683
0.8	0.0785	0.0727	0.062	0.0809	0.0743	0.0628	0.0823	0.0758	0.0653
0.9	0.0371	0.0333	0.0283	0.0372	0.0332	0.0276	0.0368	0.0331	0.0277
1.0	0.0185	0.0163	0.0129	0.0179	0.0157	0.0122	0.0170	0.0151	0.0120
1.1	0.0096	0.0081	0.0063	0.0092	0.0079	0.0060	0.0091	0.0078	0.0059
1.2	0.0052	0.0043	0.0032	0.0049	0.0041	0.0030	0.0048	0.0040	0.0029
1.3	0.0029	0.0020	0.0017	0.0027	0.0022	0.0015	0.0027	0.0022	0.0015
1.4	0.0017	0.0013	0.0009	0.0016	0.0012	0.0008	0.0016	0.0012	0.0008

TABLE B6
Critical values R_b^2 test

Model 2: $H_0 : \Delta^{d_0}(y_t - \alpha_0 - \beta_0 t) = \varepsilon_t$; $H_1 : y_t - \alpha_1 - \beta_1 t - \beta_2 DT_t = \varepsilon_t$

$d_0 T$	$T = 100$			$T = 400$			$T = 1000$		
d_0	10%	5%	1%	10%	5%	1%	10%	5%	1%
0.6	0.4059	0.3825	0.3469	0.4963	0.4715	0.4284	0.5495	0.5230	0.4779
0.7	0.1759	0.1641	0.1443	0.1959	0.1827	0.1626	0.2040	0.1906	0.1687
0.8	0.0805	0.0734	0.0627	0.0820	0.0746	0.0642	0.0834	0.0763	0.0655
0.9	0.0369	0.0331	0.0272	0.0365	0.0325	0.0272	0.0361	0.0326	0.0267
1.0	0.0153	0.0173	0.0124	0.0168	0.0147	0.0118	0.0166	0.0140	0.0112
1.1	0.0086	0.0075	0.0058	0.0083	0.0071	0.0056	0.0082	0.0070	0.0053
1.2	0.0044	0.0038	0.0028	0.0042	0.0035	0.0025	0.0040	0.0034	0.0025
1.3	0.0023	0.0019	0.0014	0.0021	0.0017	0.0012	0.0021	0.0017	0.0012
1.4	0.0012	0.0010	0.0007	0.0011	0.0009	0.0006	0.0011	0.0009	0.0006

TABLE B7

Critical values R_b^3 test

Model 3: $H_0 : \Delta^{d_0}(y_t - \alpha_0 - \beta_0 t) = \varepsilon_t$; $H_1 : y_t - \alpha_1 - \alpha_2 DC_t(\lambda) - \beta_1 t - \beta_2 DT_t = \varepsilon_t$

$d_0 T$	$T = 100$			$T = 400$			$T = 1000$		
d_0	10%	5%	1%	10%	5%	1%	10%	5%	1%
0.6	0.3654	0.3475	0.3181	0.4556	0.4352	0.4011	0.5098	0.4893	0.4513
0.7	0.1552	0.1464	0.1311	0.1764	0.1656	0.1495	0.1849	0.1743	0.1572
0.8	0.0692	0.0643	0.0563	0.0719	0.0662	0.0579	0.0735	0.0684	0.0595
0.9	0.0312	0.0290	0.0247	0.0315	0.0285	0.0244	0.0313	0.0285	0.0240
1.0	0.0150	0.0135	0.0110	0.0105	0.0126	0.0143	0.0104	0.0125	0.0140
1.1	0.0072	0.0064	0.0051	0.0070	0.0061	0.0048	0.0070	0.0060	0.0046
1.2	0.0037	0.0032	0.0025	0.0036	0.0030	0.0023	0.0034	0.0029	0.0022
1.3	0.0019	0.0016	0.0012	0.0018	0.0015	0.0011	0.0018	0.0015	0.0011
1.4	0.0010	0.0009	0.0006	0.0010	0.0008	0.0005	0.0010	0.0008	0.0006